

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-095683

(43)Date of publication of application : 08.04.1994

(51)Int.Cl.

G10L 3/00

G10L 3/00

G10L 3/00

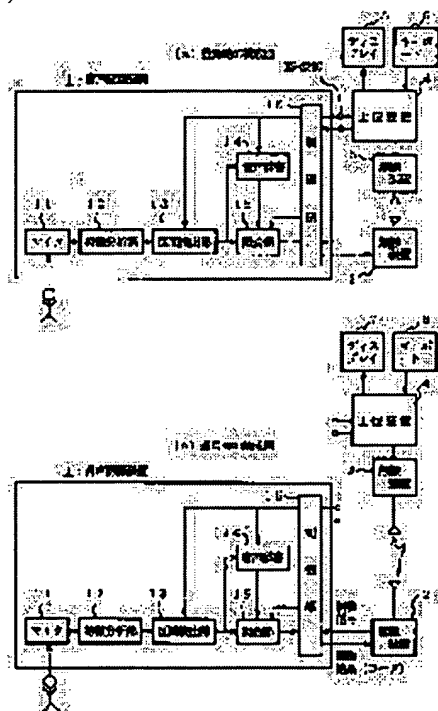
(21)Application number : 04-241570

(71)Applicant : FUJITSU LTD

(22)Date of filing : 10.09.1992

(72)Inventor : FUJIMOTO TAKAYUKI

(54) SOUND RECOGNIZING DEVICE



(57)Abstract:

PURPOSE: To connect the sound recognizing device to a radio equipment carried by an inspector, to register a sound dictionary by connecting it to a host device, to transmit a sound recognition result to the host device by radio during an operation, to prevent a reduction in a sound recognition rate caused by the deterioration of a sound quality during a transmission and to make it easier to carry it at the sound recognizing device for informing the host device of the sound recognition result by radio.

CONSTITUTION: Corresponding to the voicing of a sound to a microphone 11 by the inspector during the operation, about sound pattern data generated by analyzing a sound signal transformed into an electric signal, the word number of the sound pattern data made coincident with a sound dictionary 14 by collating it is taken out by a collating part 15, this word number is transmitted by radio by a radio device 2, it is demodulated by a radio device 3 receiving this and the word number is notified to a host device 4.

LEGAL STATUS

[Date of request for examination] 27.08.1999

[Date of sending the examiner's decision of rejection] 05.03.2002

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

THIS PAGE BLANK (USPTO)

【特許請求の範囲】

【請求項1】無線によって音声認識結果を上位装置に通知する音声認識装置において、

音声を変換するマイク（11）およびこのマイク（11）によって変換した音声信号の音声パタンデータと音声辞書（14）とを照合する照合部（15）とからなる音声認識装置（1）に無線装置（2）を接続し、

この無線装置（2）との間で無線でデータの授受を行う上位装置（4）側に無線装置（3）を備え、音声辞書（14）への登録時に、上記音声認識装置

（1）を上記上位装置（4）に接続し、検査員などが音声を上記マイク（11）に向かって発声したことに対応して、当該マイク（11）が音声を電気信号に変換し、この音声信号を分析して生成した音声パタンデータを上記音声辞書（14）に登録すると共に上位装置（4）に保存するように構成したことを特徴とする音声認識装置。

【請求項2】無線によって音声認識結果を上位装置に通知する音声認識装置において、

音声の電気信号に変換するマイク（11）およびこのマイク（11）によって変換した音声信号の音声パタンデータと音声辞書（14）とを照合する照合部（15）とからなる音声認識装置（1）に無線装置（2）を接続し、

この無線装置（2）との間で無線でデータの授受を行う上位装置（4）側に無線装置（3）を備え、音声辞書（14）の登録時に、上記音声認識装置（1）を上記上位装置（4）に接続し、音声パタンデータのダウンロード指示に対応して、上位装置（4）に保存されている音声パタンデータをダウンロードして上記音声辞書（14）に登録するように構成したことを特徴とする音声認識装置。

【請求項3】無線によって音声認識結果を上位装置に通知する音声認識装置において、

音声の電気信号に変換するマイク（11）およびこのマイク（11）によって変換した音声信号の音声パタンデータと音声辞書（14）とを照合する照合部（15）とからなる音声認識装置（1）に無線装置（2）を接続し、

この無線装置（2）との間で無線でデータの授受を行う上位装置（4）側に無線装置（3）を備え、運用時に、検査員などが上記マイク（11）に向かって音声を発声したことに対応して、電気信号に変換された音声信号を分析して生成した音声パタンデータについて、上記照合部（15）が音声辞書（14）と照合して合致する音声パタンデータの単語Noを取り出し、この単語Noを上記無線装置（2）によって無線で送信し、これを受けた上記無線装置（3）が復調して単語Noを上位装置（4）に通知するように構成したことを特徴と

する音声認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、無線によって音声認識結果を上位装置に通知する音声認識装置に関するものである。

【0002】

【従来の技術】音声認識装置の使い方として典型的なものに、図6に示すように、工場内で検査を行う検査員が当該音声認識装置から出力する音声ガイダンスを聞いて指示された検査を行い、その検査結果を音声で発声して認識させ、検査データとしてホストに通知するというシステムがある。

【0003】検査員は、検査対象のまわりを動き回りながら検査を行うため、ホストと無線を使って音声のやりとりをする。音声認識に使用する無線装置は、音声認識率を確保するため、S/N比が良くかつ周波数帯域の広いものを使うようにしていた。

【0004】以下図6の構成および動作を簡単に説明する。図6は、従来技術の説明図を示す。図6において、音声認識装置41は、無線を介して受信した検査員などの音声を認識するものであって、特徴分析部42、区間検出部43、音声辞書44、照合部45および制御部46などから構成されるものである。

【0005】特徴分析部42は、無線を介して受信した音声の特徴（エネルギーの時系列、周波数スペクトルの時系列など）を分析するものである。区間検出部43は、音の始まり、終わり、区間内のパラメタなどを検出するものである。

【0006】音声辞書44は、検査員などが予め発声した音声（単語など）の音声パタンデータを登録したものである。照合部45は、区間検出部43によって検出された情報をもとに、予め登録しておいた音声辞書44の音声パタンデータ（テンプレート）と照合し、合致するものを検出し、音声認識するものである。

【0007】制御部46は、全体を統括制御するものである。マイク47は、検査員が検査結果などを発声した音声を電気信号に変換するものである。

【0008】無線装置48は、マイク47で電気信号に変換した音声信号をアンテナから無線で送信したり、送信されてきた電波を受信して音声信号に復調したりするものである。

【0009】無線装置49は、音声認識装置41の側に設けたものであって、無線装置48から送信されてきた電波を受信して元の音声信号を取り出して音声認識装置41に入力したり、音声認識装置41からの音声信号を電波にしてアンテナから送信したりなどするものである。

【0010】上位装置50は、音声認識装置41から音声認識した認識結果を受け取り、この認識結果をもとに

所定の業務処理を行ったり、その結果を送出したりなどするものである。

【0011】次に、図6の構成の動作を説明する。

(1) 検査員がマイク47に向かって音声で部品の検査結果を発声する。マイク47はこれを音声信号に変換し、無線装置48によってアンテナから電波を放射する。

【0012】(2) (1)で放射された電波をアンテナで受信した無線装置49は、元の音声信号に復調して音声認識装置41に通知する。

(3) 音声認識装置41は、通知を受けた検査員の音声について、特徴分析部42が特徴を分析し、区間検出部43が音の始まり、終わり、区間内のパラメタの検出を行い、照合部45が音声辞書44と照合して合致する音声パターンデータを見つけ、この合致した音声パターンデータの単語N_oを認識結果として上位装置50に通知する。

【0013】(4) 上位装置50は、通知を受けた認識結果を保持すると共に確認メッセージを返答する。この確認メッセージは、音声認識装置41、無線装置49、無線装置48を通して音声や表示として検査員に知らせる。これを聞いたあるいは見た検査員は、その応答を音声でマイク47に発声し、同様にしてその音声の認識結果を上位装置50に通知する。

【0014】

【発明が解決しようとする課題】上述したように、従来は、無線装置48と無線装置49との間で音声を無線で伝送していたため、音声認識率を高めるためにS/N比が良くかつ周波数帯域の広い無線装置を使う必要があり、高価なものになってしまうという問題があった。また、電波法の規制により、同一エリアで、CH数を多く確保できないという問題もあった。特に、同一の工場内で、多くの検査員が同時に音声認識装置41を使って作業を行う場合、極めて多数のチャンネルが必要となり、これだけのチャンネル数を確保できないという問題もあった。更に、無線で音声認識を行う音声を伝送したのでは、音声の品質がどうしても劣化してしまい、音声認識率を低下させてしまうという問題もあった。

【0015】本発明は、これらの問題を解決するため、検査員などが携帯する無線装置に音声認識装置を接続すると共に音声辞書の登録を上位装置に接続して行い、運用時に音声認識結果を無線で上位装置に伝送し、伝送時の音質劣化による音声認識率の低下を防止すると共に携帯性を良好にすることを目的としている。

【0016】

【課題を解決するための手段】図1を参照して課題を解決するための手段を説明する。図1において、音声認識装置1は、音声を認識するものであって、ここでは、マイク11、音声辞書14、照合部15などから構成されるものである。

【0017】マイク11は、検査員などの音声を電気信号に変換するものである。音声辞書14は、音声パターンデータを予め登録したものである。照合部15は、音声信号の音声パターンデータと音声辞書14とを照合し、合致する音声パターンデータの単語N_oを取り出すものである。

【0018】無線装置2は、音声認識装置1と接続し、音声信号の認識結果を無線で送信したりなどするものである。無線装置3は、送信されてきた無線を受信して復元し、元の音声信号の認識結果を上位装置4に通知したりなどするものである。

【0019】上位装置4は、通知された音声認識結果をもとに各種処理を行うものである。

【0020】

【作用】本発明は、図1に示すように、音声辞書14の登録時に、音声認識装置1を上位装置4に接続し、検査員などがマイク11に向かって発声したことに対応して、当該マイク11が音声を電気信号に変換し、その音声から生成した音声パターンデータを音声辞書14に登録すると共に上位装置4に保存するようにしている。

【0021】また、音声辞書14の登録時に、音声認識装置1を上位装置4に接続し、音声パターンデータのダウンロード指示に対応して、上位装置4に保存されている音声パターンデータをダウンロードして音声辞書14に登録するようにしている。

【0022】また、運用時に、検査員などがマイク11に向かって音声を発声したことに対応して、電気信号に変換された音声信号の音声パターンデータについて、照合部15が音声辞書14と照合して合致する音声パターンデータの単語N_oを取り出し、この単語N_oを無線装置2によって無線で送信し、これを受けた無線装置3が復調して単語N_oを上位装置4に通知するようにしている。

【0023】従って、検査員などが携帯する無線装置2に音声認識装置1を接続すると共に音声辞書14の登録を上位装置4に接続して行い、運用時に音声認識結果を無線で上位装置4に伝送することにより、伝送時の音質劣化による音声認識率の低下を防止すると共に携帯性を良好にすることが可能となる。

【0024】

【実施例】次に、図1から図5を用いて本発明の実施例の構成および動作を順次詳細に説明する。

【0025】図1は、本発明の1実施例構成図を示す。図1の(a)は登録時の構成図を示し、図1の(b)は運用時の構成図を示す。

【0026】図1において、音声認識装置1は、検査員などが発声した音声を電気信号に変換して認識するものであって、ここでは、マイク11、特徴分析部12、区間検出部13、音声辞書14、照合部15および制御部16などから構成されるものである。

【0027】マイク11は、検査員などが発声した音声

を電気信号に変換するものである。特徴分析部12は、マイク11によって音声を変換した音声信号のエネルギーの時系列や周波数スペクトルの時系列などを分析するものである。

【0028】区間検出部13は、特徴分析部12によって分析された音声信号のエネルギーの時系列や周波数スペクトルの時系列から音の始まりや終わりを検出したり、区間内のパラメタを抽出したりなどするものである。

【0029】音声辞書14は、検査員などが予め発声して分析した音声パターンデータを登録した辞書である。照合部15は、音声信号を分析した音声パターンデータと、音声辞書14に格納されている音声パターンデータとを照合し、合致する音声パターンデータの単語Noを取り出すものである。

【0030】制御部16は、各種制御を行うものである。無線装置2は、無線の送受信を行うものであって、検査員などが携帯する音声認識装置1に接続した無線装置であり、検査員などが発声した音声の認識結果（ここでは、単語No）を無線で上位装置4側の無線装置3に向けて送信するものである。

【0031】無線装置3は、無線の送受信を行うものであって、上位装置4に接続し、検査員などが携帯した音声認識装置1に接続した無線装置2から送信されてきた電波を受信し、認識結果などを復調し、上位装置4に渡したりなどするものである。

【0032】上位装置4は、音声認識装置1から無線で送信されてきた音声認識結果を受け取り、各種処理を行うものである。ディスプレイ5は、各種データを表示するものである。

【0033】キーボード6は、各種データ入力を行った、各種指示を入力したりするものである。次に、図2を用いて図1の（a）の構成による音声パターンデータの登録時の動作を詳細に説明する。

【0034】図2において、S1は、作業者がキーボード6を押下して音声登録処理およびアップロード処理を上位装置4に指示する。S2は、S1の指示を有線で受信した上位装置4がプログラム起動する。これは、以降の音声パターンデータを音声辞書14に登録したり、この登録した音声パターンデータを上位装置4に転送したりなどするプログラムを起動する。

【0035】S3は、上位装置4が音声登録指示を音声認識装置1に通知する。S4は、上位装置4からS3で音声登録指示を受けた音声認識装置1が音声登録モードに移行する。

【0036】S5は、上位装置4が登録しようとする音声パターンデータの単語を表示する。これは、後述する図5の単語辞書14に示す、例えば“クラクション不良”という単語をディスプレイ5上に表示する。

【0037】S6は、作業者がディスプレイ5上の表示

を見る。S7は、作業者がS6で見た単語を発声する。例えば“ブレーキペダル不良”、“クラクション不良”などの単語をマイク11に向かって発声する。

【0038】S8は、音声認識装置1内の特徴分析部11が特徴分析を行う。この特徴分析により、音声信号からエネルギーの時系列や周波数スペクトルの時系列を生成する。

【0039】S9は、区間検出する。これは、区間検出部12が特徴分析部11によって分析されたエネルギーの時系列や周波数スペクトルの時系列から、音の始めや終わり、更に区間内のパラメタを抽出し、音声パターンデータを生成する。

【0040】S10は、音声辞書14に登録する。これは、S9で生成した音声パターンデータを音声辞書14に登録、例えば図5の音声辞書の例に示すように、単語Noに対応づけて音声パターンデータを登録する。

【0041】S11は、音声認識装置1が登録終了通知を上位装置4に行う。S12は、S11で登録終了通知を受けた上位装置4が音声登録の終了か判断する。YESの場合には、音声登録すべき全て単語を表示してその単語パターンデータを音声辞書14に登録したので、S13に進む。一方、NOの場合には、次の音声登録指示をS3で行うことを繰り返す。

【0042】S13は、音声辞書の転送指示を音声認識装置1に通知する。これは、音声登録を終了したので、音声辞書14の内容（音声パターンデータ）を上位装置4に転送する旨の指示を行う。

【0043】S14は、S13で転送指示を受けた音声認識装置1が音声辞書14から音声パターンデータを取り出す。S15は、音声認識装置1が音声辞書14から取り出した音声パターンデータを上位装置4に送出する。

【0044】S16は、S15で送出されてきた音声パターンデータを保存し、ダウンロードに備える。以上によって、作業者が携帯する音声認識装置1を上位装置4にRS-232Cによって直接に接続し、当該上位装置4のキーボード6やディスプレイ5を使ってこれからの作業に必要な音声パターンデータを音声辞書14に登録すると共に、上位装置4に保存する。これにより、作業者が携帯する音声認識装置1および無線装置2の側に、かさばるディスプレイ5やキーボード6を装着する必要がなくなり、小型軽量にすることが可能となる。また、音声辞書14に登録した音声パターンデータを上位装置4に転送して保存し、後述する図3で説明するように、上位装置4から音声認識装置1の音声辞書14にダウンロードすることにより、その都度、作業者が音声を発声してその音声パターンデータを生成して音声辞書14に登録する手間が不要となり、迅速に音声辞書14に登録することが可能となる。

【0045】図3は、図1の（a）の構成における音声辞書ダウンロード処理の動作説明図を示す。これは、図

2のS16で上位装置4に保存した音声パタンデータを音声認識装置1の音声辞書14にダウンロードするときの動作説明図である。

【0046】図3において、S21は、作業者がキーボードを押下し、音声辞書ダウンロード指示を行う。S22は、S21でダウンロード指示を受けた上位装置4がプログラム起動する。

【0047】S23は、上位装置4が音声辞書ダウンロード指示を音声認識装置1に通知する。S24は、上位装置4が音声辞書ダウンロードを行う。これは、図2のS16で保存しておいた音声辞書14の音声パタンデータを取り出し、RS-232Cを通して音声認識装置1にダウンロード（転送）する。

【0048】S25は、音声認識装置1がダウンロードされてきた音声パタンデータを音声辞書14に登録する。以上によって、音声認識装置1の音声辞書14に対して、上位装置4に保存しておいた音声パタンデータを登録することにより、その都度、作業者が発声して生成した音声パタンデータを音声辞書14に登録する手間がなくなり、迅速に音声パタンデータを音声辞書14に登録することが可能となる。

【0049】次に、図4を用いて図1の（b）の構成による運用時の動作を詳細に説明する。図4において、S31は、作業者がキーボード6を押下して音声認識処理を開始する旨を上位装置4に通知する。

【0050】S32は、S31の指示を受けた上位装置4がプログラム起動する。これは、以降の運用を行うプログラムを起動する。S33は、上位装置4が音声認識指示を無線で音声認識装置1に通知する。

【0051】S34は、上位装置4からS33で音声認識指示を受けて音声認識装置1が音声認識モードに移行する。S35は、作業者が音声認識装置1およびこれと一体となっている無線装置3を携帯して作業現場に移動する。

【0052】S36は、作業者がここでは、検査のチェック結果を発声する。例えば“ブレーキペダル不良”や“クラクション不良”などをマイク11に向けて発声する。S37は、音声認識装置1内の特徴分析部11が特徴分析を行う。この特徴分析により、音声信号からエネルギーの時系列や周波数スペクトルの時系列を生成する。

【0053】S38は、区間検出する。これは、区間検出部12が特徴分析部11によって分析されたエネルギーの時系列や周波数スペクトルの時系列から、音の始まりと終わりを検出、更に区間内のパラメタを抽出し、音声パタンデータを生成する。

【0054】S39は、照合処理15がS38で生成した音声パタンデータと、音声辞書14内の音声パタンデータとを照合する。S40は、選ばれた音声パタンデータの単語Noを取り出す。これは、S39で音声信号から生成した音声パタンデータと、音声辞書14の音声パ

タンデータとを照合し、合致した音声パタンデータの単語Noとして、例えば図5の音声辞書14から“ブレーキペダル不良”に対応する単語No“n+3”を取り出す。

【0055】S41は、認識結果（S40で取り出した単語No）を上位装置4に無線で通知する。S42は、上位装置4は音声の認識結果（ここでは、単語No）を保持する。

【0056】S43は、上位装置4が無線で確認メッセージ、例えば“クラクション不良”を通知する。この確認メッセージの通知は、無線で音声認識装置1に通知してスピーカあるいはイヤホンで作業者に知らせたり、あるいはディスプレイ上に表示して知らせる。

【0057】S44は、作業者がS43の確認メッセージを聞いたりあるいは見たりし、合っていれば“YES”とマイク11に向かって発声し、間違っていれば“NO”とマイク11に向かって発声する。

【0058】S45は、上記と同じ認識処理（S37からS40）を行う。S46は、認識結果（ここでは単語Noのコード）を上位装置4に通知する。S47は、S46で認識結果の通知を受けた上位装置4が“YES”以外ならば前回の認識結果データを破棄する。これにより、作業者が音声で発声して自動認識した単語Noを受信した上位装置4が確認メッセージ（音声あるいは表示）して作業者から“YES”を受信したときはその認識結果データを有効として業務処理に使用し、一方、“NO”を受信したときはその認識結果データを破棄する。

これにより、作業者は、音声認識装置1および無線装置2が一体となった小型な装置を携帯し、そのマイク11に向かって検査結果を発声すると、音声でその確認メッセージを聞いたり、表示パネルに表示された内容を見て“YES”あるいは“NO”を発声して確認することが可能となる。従って、作業者は、携帯した音声認識装置1内で音声を認識した認識結果（ここでは単語No）のみをコードで上位装置4に無線で送信するのみでよく、従来の音声自体を無線で上位装置4に送信してこれを音声認識するシステムに比し、無線で伝送するデータ量を削減でき、同一工場内で同じ周波数帯域内で非常に多くの作業者が同時に無線を使って検査業務などを行うことが可能となる。この際、音声認識装置1の音声辞書14に音声パタンデータを登録する際に、上位装置4にRS-232Cなどで直接接続して登録を行うと共に予め登録した音声パタンデータをダウンロードしたりし（図1の（a）の構成、図2、図3参照）、運用時に無線で認識結果のみを上位装置4に送信し、小型軽量にすることができる。

【0059】図5は、本発明の音声辞書例を示す。ここで単語Noは、一意に付与した音声辞書データの番号である。音声辞書データは、単語の音声パタンデータであって、音声を認識するための周波数スペクトルの時系列

のデータであり、予め作業者が発声してそのときの周波数スペクトルの時系列のデータを登録したものである。

【0060】ここでは、“1”、“2”などの作業者が発声した音声の音声パタンデータを単語Noに対応づけて図示のように予め登録する。同様に、単語“YES”、“NO”、更に“クラクション不良”、“ブレーキペダル不良”などと発声した音声パタンデータを図示のように登録する。

【0061】このようにして予め単語の音声パタンデータを音声辞書14に登録しておき、作業者が作業中に音声でマイク1に向かって発声したときにこの音声の音声パタンデータを生成し、当該音声辞書14内に登録されている音声パタンデータとパタンマッチングを行い、最も合致する割合の高いものの単語Noを音声認識した結果として取り出す。この取り出した認識結果(単語No)を無線によって上位装置4に通知する。

【0062】

【発明の効果】以上説明したように、本発明によれば、検査員などが携帯する無線装置2と音声認識装置1を一体とすると共に音声辞書14の登録を上位装置4に接続して行い、運用時に音声認識結果を無線で上位装置4に伝送する構成を採用しているため、伝送時の音質劣化による音声認識率の低下を防止し、データ量を少なくして信頼性が高い音声の認識結果を上位装置4に通知することができる。これらにより、

(1) 同一エリア内で従来に比し、CH数を多く確保でき、同一工場内で多くの検査員が携帯した音声認識装置1を同時に使って作業を行うことができる。

【0063】(2) 従来の無線で音声を送信していた場合に発声した音声品質の劣化による音声認識率の低下を、本発明では防止でき、音声認識率が高くなる。

(3) 従来の無線で音声を送信するときに要求されていた音声の品質を良好にする必要がなくなり、認識結果をコードなどで伝送すればよく、高価な無線機を使わなくてもよくなり、コストを低減および小型化できる。

【図面の簡単な説明】

【図1】本発明の1実施例構成図である。

【図2】図1の(a)の動作説明図(音声登録処理、アップロード処理)である。

【図3】図1の(a)の動作説明図(音声辞書ダウンロード処理)である。

【図4】図1の(b)の動作説明図(音声認識処理)である。

【図5】本発明の音声辞書例である。

【図6】従来技術の説明図である。

【符号の説明】

1：音声認識装置

11：マイク

12：特徴分析部

13：区間検出部

14：音声辞書

15：照合部

16：制御部

2、3：無線装置

4：上位装置

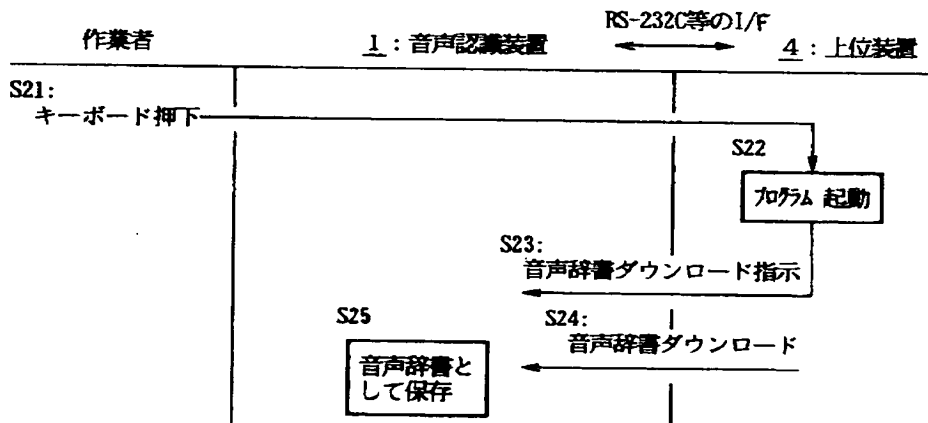
5：ディスプレイ

6：キーボード

【図3】

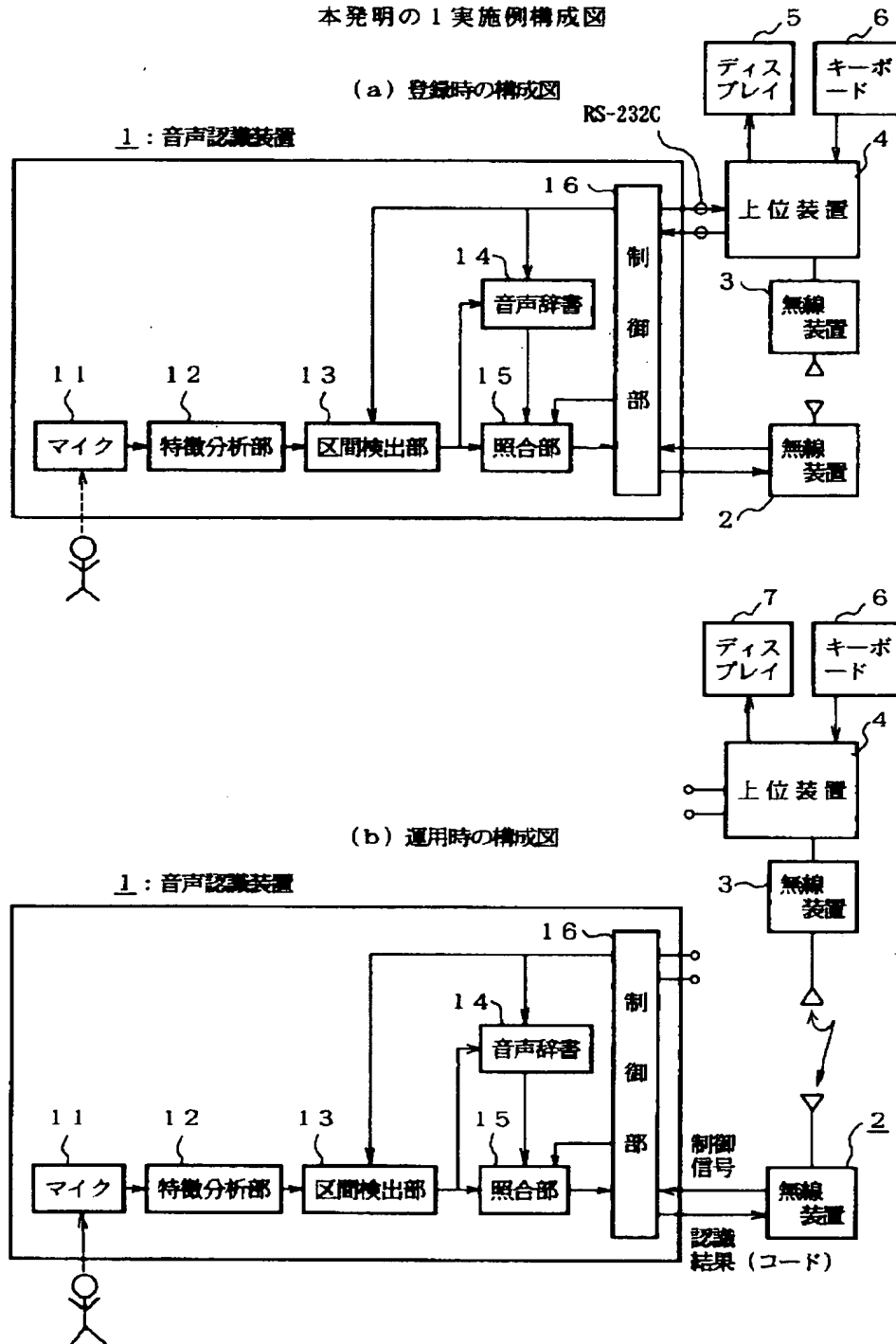
図1の(a)の動作説明図

(音声辞書ダウンロード処理)



【図1】

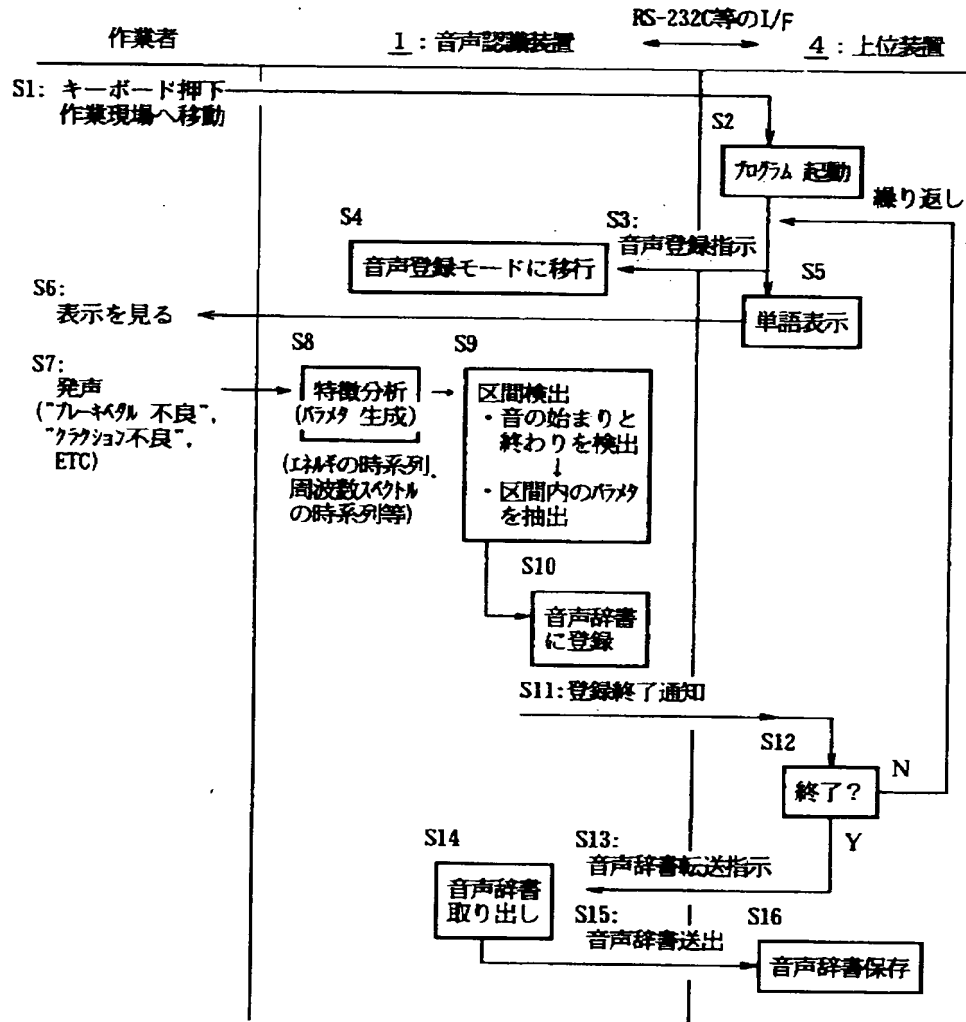
本発明の1実施例構成図



【図2】

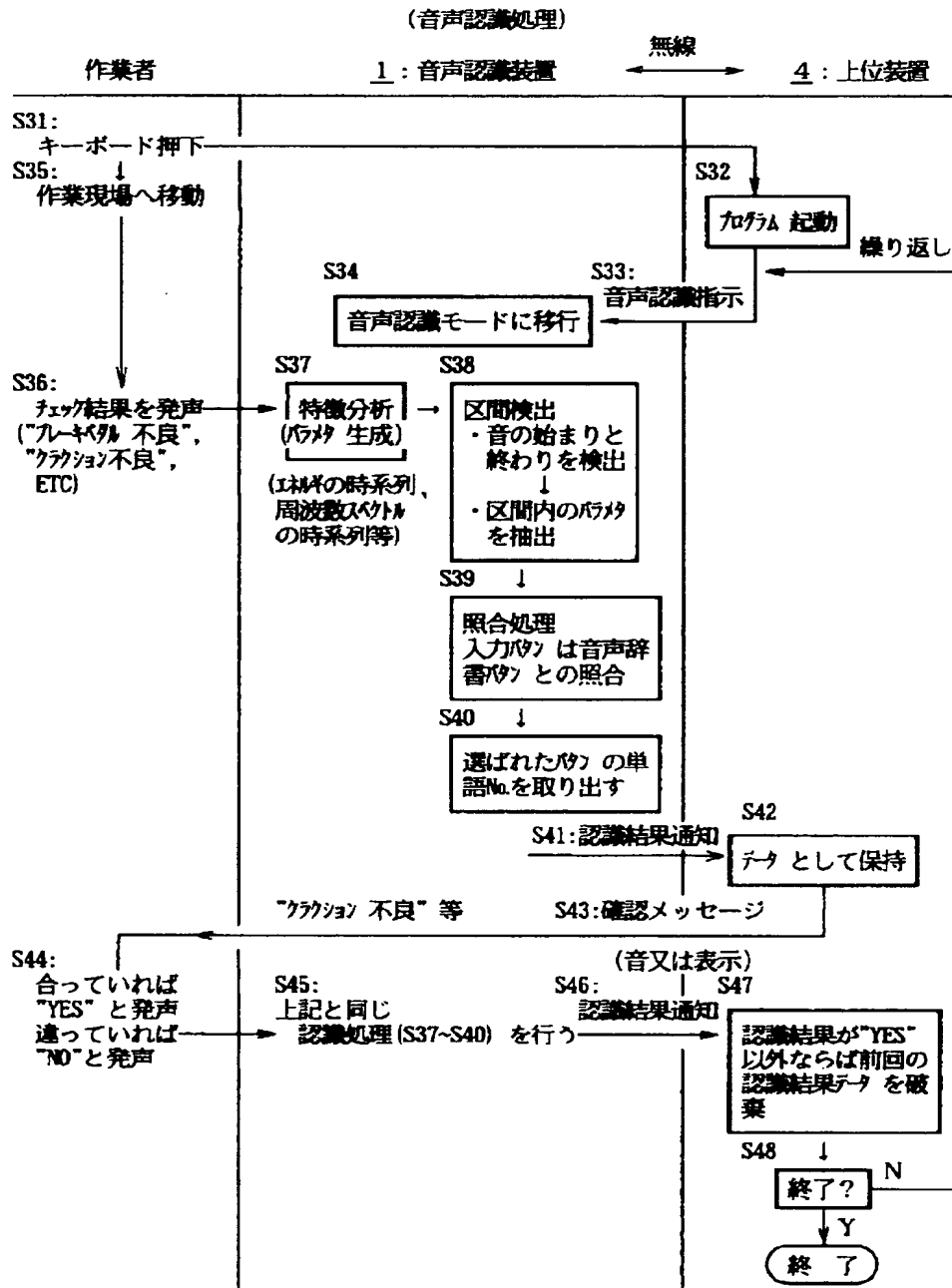
図1の(a)の動作説明図

(音声登録処理, アップロード処理)



【図4】

図1の(b)の動作説明図



【図5】

本発明の音声辞書例

音声辞書の例

単語No.	音声辞書データ
1	" 1 " の音声ボタンデータ
2	" 2 " の音声ボタンデータ
3	.
⋮	⋮
⋮	⋮
n	" YES " の音声ボタンデータ
n+1	" NO " の音声ボタンデータ
n+2	" クラクション不良 " の音声ボタンデータ
n+3	" ブレーキペダル不良 " の音声ボタンデータ
⋮	⋮
⋮	⋮
N	" チェック終了 " の音声ボタンデータ

【図6】

従来技術の説明図

